

周报

冯浩哲

2018.12.16

周报

Work

工作进度

论文阅读

Work

在LIDC-IDRI数据集上对VAE进行了训练，同时对VAE对于相似度度量的结果进行了可视化。

工作进度

项目	进度	截止时间
无监督推荐标注	在肺结节数据集上完成了变分自编码器的训练，开始用VAE得到的KL散度进行kNN分类与推荐。我们用KL散度进行度量，发现KL散度确实可以对相似度进行很好的刻画，同时用KL散度进行kNN可以达到81.3%的肺结节分类准确率。	
CAD学报论文	在编辑的指导下完成了论文的格式修改	

论文阅读

1. Markov Chain Monte Carlo and Variational Inference: Bridging the gap

该文献是VAE的作者Kingma的论文，它提出了一种算法，该算法可以将MCMC采样方法与变分推断结合。

注意到MCMC方法是一种基于Markov Chain的Monte Carlo随机采样方法，只要采样次数够多，它就能足够逼近任意后验分布，MCMC方法用 $z_t \sim q(z_t|z_{t-1}, x)$ 来进行采样。

同时，变分推断的目的是最大化变分下界，即：

$$\log(p(x)) \geq \log(p(x)) - D_{KL}(q_\theta(z|x)||p(z|x)) = E_{q_\theta(z|x)}[\log(p(x, z)) - \log(q_\theta(z|x))] \quad (1)$$

注意当我们进行预测的时候，我们一般对(1)式给出的是它的一个无偏估计，也就是采样 $z' \sim q_\theta(z|x)$ ，然后计算 $\log(p(x, z')) - \log(q_\theta(z'|x))$ 作为下界的估计，再最大化它就可以了。

当然我们也可以进行多次预测，然后取一个均值等等。

这篇文章要解决的问题是：

我能否用MCMC方法对 z 进行采样，然后假设我采样了 T 个，比如 $z' = (z_0, z_1, \dots, z_T)$ ，那我如何用这个MCMC来对(1)的下界进行优化呢？

文中用的比较核心的东西是：

$$q(z|x) = q(z_0|x) \prod_{t=1}^T q(z_t|z_{t-1}, x) \quad (2)$$

作为其边界，同时有：

$$\begin{aligned} y &= (z_0, \dots, z_{t-1}) \\ E_{q_\theta(z|x)}[\log(p(x, z)) - \log(q_\theta(z|x))] &= E_{q_\theta(y, z_T|x)}[\log(p(x, z_T, y)) - \log(q_\theta(y, z_T|x))] \end{aligned} \quad (3)$$

然后引入辅助分布：

$$p(x, z_T, y) = p(x, z_T)r(y|x, z_T) \quad (4)$$

这样就把MCMC方法和变分推断结合起来了，从而给出了变分下界更好的估计，并让预测更加准确。